



# Australian Bureau of Statistics

## 1200.0.55.005 - Language Standards, 2012, Version 1.1

Previous ISSUE Released at 11:30 AM (CANBERRA TIME) 26/09/2012 First Issue

---

## Summary

### Contents

#### Contents

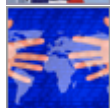
---



#### First Language Spoken



#### Languages Spoken at Home



#### Main Language Other Than English Spoken at Home



#### Main Language Spoken at Home



#### Proficiency in Spoken English

### References

## References

---

- ABS (Australian Bureau of Statistics), 2011. Australian Standard Classification of Languages (ASCL), 2011 (cat. no. 1267.0), ABS, Canberra.
- ABS, 2011, 2011 Census Dictionary, (cat. no. 2901.0), ABS, Canberra.
- ABS, 1999, Standards for Statistics on Cultural and Language Diversity, (cat. no. 1289.0), ABS, Canberra.
- Department of Employment, Education and Training, 1990, Australia's language: the Australian language and literacy policy: companion volume to the policy information paper, Australian Government Publishing Service. published on the web: <http://www.voced.edu.au/content/ngv37015>.
- Department of Employment, Education and Training, 1990. The Language of Australia: Discussion Paper on an Australian Literacy and Language Policy for the 1990's, Australian Government Publishing Service.
- Department of Immigration and Border protection, 2013, <http://www.immi.gov.au/skills/skillselect/index/functional-english/>.
- Ingram, D E, 2003, English Language Policy in Australia: Paper to the 2003 Summer International Conference of the Korea Association of Teachers of English (KATE), Chungnam National University, Daejeon City, Korea
- Ministerial Council for Education, Early Childhood Development and Youth Affairs (MCEECDYA), 2009, *Aboriginal and Torres Strait Islander Education Action Plan 2010 – 2014*.
- SCRGSP (Steering Committee for the Review of Government Service Provision) 2009, *Overcoming Indigenous Disadvantage: Key Indicators 2009*, Productivity Commission, Canberra. Ministerial Council for Education, Early Childhood Development and Youth Affairs (MCEECDYA), 2009, *Aboriginal and Torres Strait Islander Education Action Plan 2010 – 2014*.
- Stromback, T., & Preston, A., 1990, The Cost of Low Levels of English Proficiency

## Proficiency in Spoken English

### Proficiency in Spoken English

---

The Proficiency in Spoken English Statistical Standard has recently been reviewed by the ABS.

The revised standard can be accessed by clicking the Proficiency in Spoken English link on the contents page.

## About this Release

The language standard variables are used within and outside the ABS to measure language diversity and usage in Australia. They include the variables; First Language Spoken, Language Spoken at Home, Main Language Other than English Spoken at Home, and Main Language Spoken at Home. These variables provide a means for standardising the way language data are collected and disseminated.

## Expanded Contents

### CONTENTS

#### First Language Spoken

- Introduction - FLS

- Underlying Concepts - FLS

- Classification and Coding - FLS

- Collection Methods - FLS

- Outputs - FLS

- Measurement Issues - FLS

- Appendix A - Explanatory Script - FLS

#### Languages Spoken at Home (LSH)

- Introduction - LSH

- Underlying Concepts - LSH

- Classification and Coding - LSH

- Collection Methods - LSH

- Outputs - LSH

- Measurement Issues - LSH

- Appendix A - Explanatory Script - LSH

#### Main Language Other Than English Spoken at Home (MLOTESH)

- Introduction - MLOTESH
- Underlying Concepts - MLOTESH
- Classification and Coding - MLOTESH
- Collection Methods - MLOTESH
- Outputs - MLOTESH
- Measurement Issues - MLOTESH
- Appendix A - Explanatory Script - MLOTESH

Main Language Spoken at Home (MLSH)

- Introduction - MLSH
- Underlying Concepts - MLSH
- Classification and Coding - MLSH
- Collection Methods - MLSH
- Output - MLSH
- Measurement Issues - MLSH
- Appendix A - Explanatory Script - MLSH

Proficiency in Spoken English

- Introduction - PSE
- Underlying Concepts - PSE
- Classification and Coding - PSE
- Collection Methods - PSE
  - Collection Methods (PSE) - Filter Questions Modules for Standard Question
  - Collection Methods (PSE) - Standard Question Modules
  - Collection Methods (PSE) - Short Question Modules
- Output PSE
- Measurement Issues - PSE
- Appendix A - Explanatory Script - PSE
- Appendix B - Explanatory Notes - Migrant English Language Testing

## First Language Spoken

[Contents >> First Language Spoken](#)

This section contains the following subsection :

- Introduction - FLS
- Underlying Concepts - FLS
- Classification and Coding - FLS
- Collection Methods - FLS
- Outputs - FLS
- Measurement Issues - FLS
- Appendix A - Explanatory Script - FLS

[Previous Page](#)

[Next Page](#)

## Introduction - FLS

[Contents >> First Language Spoken >> Introduction - FLS](#)

## First language Spoken - Introduction

---

There is a demand for objective data about languages in Australia. Drivers for these data include government policy on language and literacy, which has evolved rapidly over recent years. Key to the policy are English literacy and the economic relevance of language skills, combined with the significance of language to multiculturalism and diversity within Australia. The goals of Australia's language policy emphasise:

- the importance for all Australians to maintain an appropriate level of written and spoken English
- the importance of Australians learning languages other than English
- the maintenance and learning of Aboriginal and Torres Strait Islander languages reinforced by the Council of Australian Governments' (COAG) policy 'Closing the Gap' and
- the provision of translating and interpreting services.

Stakeholders for language data include health and education service providers, data collection agencies, government agencies, community groups, academics and private sector organisations. Stakeholder interests relate to methodologies for collecting and classifying language information.

To meet these data needs, the Australian Bureau of Statistics (ABS) has incorporated language questions into a range of social statistical collections. ABS has developed standards for a number of language variables for use when collecting language data both within and outside the ABS.

The purpose of this First Language Spoken variable is to measure the language a person first spoke rather than the language they could first read or write. To ensure that all relevant languages can be measured, the variable includes sign languages.

Program providers have identified that disadvantage in access and equity to government and community programs and services may be experienced by people in Australia whose first language is not English. Data relating to First Language Spoken may aid as an indicator of deficiency in English language competency and other factors associated with cultural background.

The variable First Language Spoken is one of the standard set of indicators for use in measuring cultural and linguistic diversity. It can be used in conjunction with others in the set to understand cultural and ethnic diversity within Australia. Its implementation in administrative and service settings enables the provision of relevant data to assess, measure and monitor service needs, including access and equity requirements.

## First Language Spoken - Underlying Concepts

---

### Name of variable

The name for the variable is First Language Spoken.

### Definition of variable

#### Nominal definition

First Language Spoken is defined as the first language a person masters during the language acquisition phase of intellectual development. This would generally be the language spoken in the home by the people who have raised the person from infancy.

First Language Spoken is an attribute of the counting unit 'person'.

#### Operational definition

Operationally, First Language Spoken is defined as the language the respondent identifies, or remembers, as being the first language which they could understand to the extent of being able to conduct a conversation.

The definition of language is provided in the Australian Standard Classification of Languages (ASCL), 2011 (ABS cat. no. 1267.0)

### Discussion of conceptual issues

The concept of first language includes the terms 'native language' and 'mother tongue.' However, within the Australian context 'First Language Spoken' is the appropriate term. The term 'first language' is used in Australian government policy documents and referred to in linguistic journals.

The 'First Language Spoken' variable and the ASCL recognise that approximately one percent of the Australian population use non verbal forms of communication. For coding purposes Auslan and similar sign languages are recognised as separate languages. However Signed English/finger spelling is considered to be another form of English and coded against English.

The question module is, 'Which language did you first **speak** as a child?' ABS question testing indicates use of the word 'speak' poses few response problems for people who use sign language.

The question modules used for this variable are not designed to measure dual first language speakers and should not be used for this purpose. Although there is some academic interest in the identification of respondents who claim to have spoken two languages as first languages (for example in the study of changing language patterns, cultural affiliation or expected language proficiency), research indicates that data produced from a question seeking this information may not be statistically viable.

First Language Spoken is one of five language variables. The other language variables are Languages Spoke at Home, Main Language Spoken at Home, Main language other Than

## Classification and Coding - FLS

[Contents](#) >> [First Language Spoken](#) >> [Classification and Coding - FLS](#)

### First Language Spoken - Classification and Coding

---

#### The standard classification criteria and structure

Language data in Australia should be collected, aggregated and disseminated using the Australian Standard Classification of Languages (ASCL), 2011 (ABS cat. no. 1267.0). The two classification criteria used to form categories in the ASCL are:

- genetic affinity, which is the relationship between languages due to evolution from a common ancestral language and
- geographic proximity of language(s) based on the area in which the language originated or first became recognised as a distinct language.

The classification has a three level hierarchical structure, with the exception of the Australian Indigenous Languages broad group, where an extra level has been added between the narrow group and language levels of the classification.

The most detailed level of the classification consists of 432 (216 Indigenous and 216 non Indigenous) base level units. Included in the 432 base level units are 388 specific languages and 44 'not elsewhere classified' (nec) categories, used to code languages not separately listed in the classification.

The second level of the classification comprises 51 narrow groups of languages that are similar in terms of the classification criteria, including seven 'other' categories which consist of languages which do not fit into a particular narrow group.

For three narrow groups of Australian Indigenous languages (Narrow Group 81 Arnhem Land and Daly River Region Languages, Narrow Group 82 Yolngu Matha and Narrow Group 86 Arandic) three digit levels are positioned between the narrow group and language level of the classification. There are 13 three digit level categories. They provide meaningful and useful groups of languages.

The first and most general level of the classification comprises nine broad groups of languages including one 'other' category. Broad groups are formed by aggregating geographically proximate narrow groups.

## The code structure

One, two and four digit codes are assigned to the first, second and third level units of the classification respectively. The first digit identifies the broad group in which each language or narrow group is contained. The first two digits taken together identify the Narrow Group in which each Language is contained.

Within Broad Group 8, Australian Indigenous Languages, there are some narrow groups where an extra level has been added between the narrow group and language levels of the classification. The first three digits taken together identify the additional language groups at this third level.

The four digit codes represent each of the 432 Languages or base level units.

The following examples illustrate the coding scheme:

- |    |              |      |                                 |
|----|--------------|------|---------------------------------|
| a. | Broad Group  | 5    | Southern Asian Languages        |
|    | Narrow Group | 52   | Indo-Aryan                      |
|    | Language     | 5206 | Nepali                          |
|    |              |      |                                 |
| b. | Broad Group  | 8    | Australian Indigenous Languages |
|    | Narrow Group | 82   | Yolngu Matha                    |
|    | Extra Level  | 824  | Dhuwala                         |
|    | Language     | 8245 | Madarrpa                        |

## Residual categories and codes

Within each narrow group, a four digit code, consisting of the narrow group code followed by the digits '99', is reserved for a residual 'not elsewhere classified' (nec) or 'other' category. Similarly, within the third level classification of Australian Indigenous Languages the three digit group code may be followed by '9' to denote a 'not elsewhere classified' (nec) or 'other' category. All languages which are not separately identified in the classification are included in these residual 'nec' or 'other' categories of the related classification level.

In each broad group, two digit codes are reserved for residual categories at the narrow group level. These codes consist of the broad group code followed by '9'. These categories are termed 'Other *Broad Group Name*' and consist of separately identified languages which do not fit into any of the narrow groups contained within the broad group, based on the classification criteria.

## Supplementary codes

Supplementary codes are not part of the classification structure. They exist for operational reasons only, and no data would be coded to them if sufficiently detailed responses were obtained in all instances. They are used to process inadequately described responses in statistical collections. The codes are of two types:

- four digit codes ending with one, two or three zeros
- four digit codes commencing with three zeros (operational codes).

Codes ending in zero are described as 'not further defined' (nfd) codes. These codes classify responses to a question about language which cannot be coded to the language



level of the classification, but which can be coded to a higher level of the classification structure.

Responses which do not relate directly to a particular language category, but which are within the range of languages relating to a particular narrow group, are coded to that narrow group. Such responses are allocated an 'nfd' code consisting of the two digit code of the narrow group followed by 00.

Language responses which do not directly relate to a particular narrow group or language category, but are within the range of languages relating to a particular broad group, are coded to that broad group. These responses are allocated an 'nfd' code consisting of the one digit code of the broad group followed by 000. Language responses which can only be coded at the broad or narrow group levels of the classification can be processed within a collection coded at the four digit level.

Four digit codes commencing with 000 are supplementary codes included for operational purposes to facilitate the coding of responses such as inadequately described languages, etc., which contain insufficient information to be allocated a language, narrow group or broad group code.

### **Scope of the variable**

The variable First Language Spoken applies to all persons.

### **Application of the classification to other variables**

The ASCL can be used for a variety of variables. These include: Main Language Spoken at Home, Main Language Other Than English Spoken at Home, Languages Spoken at Home, Language of Greatest Competency, and Preferred Language.

### **Coding procedures and coding Index**

Language responses to the First Language Spoken question are coded to the ASCL, or to one of the supplementary codes, using the guidelines detailed in that classification.

A coding index has been developed to assist in the implementation and use of the ASCL. It contains a comprehensive list of probable responses to questions relating to language and their correct classification codes. Each language response is matched with an entry in the ASCL Coding Index to determine the correct code. Use of the coding index enables responses to be coded accurately to the appropriate category of the classification.

Further information about the classification criteria and coding of data about languages can be found in the ASCL.

Copies of the [Coding Indexes](#) can be found in the ASCL data cube on the ABS website ([www.abs.gov.au](http://www.abs.gov.au)).

## Collection Methods - FLS

Contents >> First Language Spoken >> Collection Methods - FLS

### First Language Spoken - Collection Methods

---

#### Standard Question Modules

There are two standard question modules for First Language Spoken:

- The Question Module for Detailed Data elicits more comprehensive responses and data on First Language Spoken.

There are two alternatives for the Question Module for Detailed Data which may be used, depending on space and cost considerations:

- Alternative one: a set of 'most likely' languages with tick boxes and an 'Other - please specify' response category for writing a language not specified in the tick box list
- Alternative two: a tick box for 'English' and an 'Other - please specify' write-in category

Alternative two is more compact but will incur additional coding costs.

- The Question Module for Minimum Data contains only the tick box options 'English' and 'Other' and therefore does not collect details as to which language other than English is spoken.

#### Question module for detailed data

##### *Alternative one*

This standard question module for detailed data for First Language Spoken is recommended for all collections designed to collect extensive data on first language spoken.

The standard question module for obtaining detailed data for First Language Spoken follows:

**Q1. Which language [did you ] [did the person] [did (name)] [will (name of child under two years)] first speak as a child?**

English	q
Mandarin	q
Italian	q
Arabic	q
Cantonese	q
Greek	q
Vietnamese	q

Spanish	q
Hindi	q
Tagalog	q
Other - Please Specify.....	

For self enumerated questionnaires, respondents should be instructed to mark one box only.

The tick box list for this question module includes languages based on their statistical significance in the Australian context. The list is designed for efficiency when coding responses. The list is reviewed when data indicate that different languages have been more or less frequently reported in the ABS Census.

The tick box list can be extended where circumstances enable longer lists, consistent with the ASCL, to be displayed, such as in electronic collection drop down lists. Similarly, if there are space constraints on paper forms, the tick box list can be truncated.

The 'Other - please specify' category is included to enable respondents, whose first language spoken is not listed, to write their first language in the a space provided.

### ***Alternative two***

This alternative question may be used when detailed language data is required from a survey instrument which has space constraints. The question uses only the tick box for 'English' and the 'Other - please specify' option for respondents to write in their language. Additional coding costs will be incurred since each response will need to be matched to an entry in the Australian Standard Classification of Languages (ASCL), 2011 (ABS cat. no. 1267.0).

**Q1. Which language [did you ] [did the person] [did (name)] [will (name of child under two years)] first speak as a child?**

English	q
Other - Please Specify.....	

### **Question module for minimum data**

The standard question module for minimum data for the variable First Language Spoken follows:

**Q1. Which language [did you ] [did the person] [did (name)] [will (name of child under two years)] first speak as a child?**

English	q
Other	q

This minimum data question is suitable for administrative collections or surveys where there is no requirement for detailed language data. It can be used to identify respondents who may require assistance to access services and information, due to inadequate English language skills.

Each alternative to the question module may be accompanied by a brief explanatory note about why Language data is collected and detailed instructions about how to answer the question. The explanatory note can be included with the chosen question module or in supplementary documentation. The recommended text for the explanatory note is in

## **Standard input categories**

The standard input categories for First Language Spoken are the same as the categories of the ASCL plus the supplementary codes.

## **Input categories for minimum data**

The standard input categories for the Question Module for Minimum Data are:

- English
- Other language.

[Previous Page](#)

[Next Page](#)

# **Outputs - FLS**

[Contents >> First Language Spoken >> Outputs - FLS](#)

## **First Language Spoken - Output**

---

## **Standard output categories**

### **Output categories for detailed data**

The hierarchical structure of the Australian Standard Classification of Languages (ASCL), Second Edition (ABS cat. no. 1267.0) allows users the flexibility to output statistics at the level of the classification which best suits their requirements. Data can be presented at broad group level, narrow group level, or the base (language) level. If necessary, significant languages within a narrow group can be presented separately while the remaining languages within the narrow group are aggregated. This approach can also be applied to highlight significant narrow groups within a broad group.

Regardless of the level of aggregation intended for the dissemination of statistics, data should be captured, classified and stored at the base-level of the classification wherever possible. Collecting and storing data at the most detailed level of the classification allows the greatest flexibility for the output of statistics, enables more detailed and complex analyses, facilitates comparisons with previous data using different classifications, and preserves information that provides maximum flexibility for future use of the data.

However, the constraints affecting each statistical collection or other application, such as problems with confidentiality or standard errors, may affect the collection and/or output of data at the lower levels of the classification.

## Output categories for minimum data

The standard output categories for the Question Module for Minimum Data are:

- English
- Other language.

## Supporting variables

The variable First Language Spoken requires no supporting variables.

[Previous Page](#)

[Next Page](#)

# Measurement Issues - FLS

[Contents](#) >> [First Language Spoken](#) >> [Measurement Issues - FLS](#)

## First Language Spoken - Measurement Issues and Related Classifications

---

### Measurement issues

No specific measurement issues apply to the variable First Language Spoken beyond the normal constraints affecting statistical collections, such as problems with confidentiality or standard errors which may affect the collection or output of data for some of the categories at the lower levels of the classification.

### Related classifications

Not applicable.

### Contact information

The First Language Spoken statistical standard has been produced by the Australian Bureau of Statistics. Enquiries should be directed to [standards@abs.gov.au](mailto:standards@abs.gov.au).

[Previous Page](#)

[Next Page](#)

# Appendix A - Explanatory Script - FLS

[Contents](#) >> [First Language Spoken](#) >> [Appendix A - Explanatory Script - FLS](#)

## Appendix A - Explanatory Script - First Language Spoken

---

**Why are these question modules asked?**

The variable First Language Spoken is one of the standard set of indicators used in measuring cultural and linguistic diversity. Language data related to First Language Spoken can be used as an indicator of disadvantage, which may be experienced by people whose first language is not English.

### **Question module for Detailed Data**

#### **Alternative One**

Q.1. Which language [did you] [did the person] [did (name)] [will (name of child under two years)] first speak as a child?

English

Mandarin

Italian

Arabic

Cantonese

Greek

Vietnamese

Spanish

Hindi

Tagalog

Other - please specify .....

#### **Interview based questionnaires**

An interviewer marks the tick box for a language the respondent identified as first spoken. If the first language spoken is not a tick box option, the interviewer writes it in the 'Other - please specify' category. If a respondent identifies that they first spoke English and another language, the interviewer can prompt for the predominant first language spoken.

#### **Self enumerated questionnaires**

The respondent marks one box for their first language spoken in the tick box list. If a respondent's first language spoken is not a tick box option, they write one language in the 'Other - please specify' category.

For people who cannot speak, 'not able to speak' can be written in the 'Other - please specify' category.

#### **Alternative Two**

Q.1. Which language [did you] [did the person] [did (name) ] [will (name of child under two years)] first speak as a child?

English

Other - please specify .....

#### **Interview based questionnaires**

An interviewer marks the 'English' box for respondents who identify English as their first language spoken. If a respondent identifies another language as their first language spoken, the interviewer writes the language in the 'Other - please specify' category.

#### **Self enumerated questionnaires**

The respondent marks the 'English' box if it is their first language spoken. If English is not identified as their first language spoken, the respondent writes one language only in the 'Other - please specify' category.

For people who cannot speak, 'not able to speak' can be written in the 'Other - please specify' category.

### **Question module for Minimum Data**

Q.1. Which language [did you] [did the person] [did (name)] [will (name of child under two years)] first speak as a child?

English

Other

### **Interview based questionnaires**

An interviewer marks the 'English' box for respondents who identify English as their first language spoken. If a respondent identifies another language, the interviewer puts a mark in the 'Other' box.

### **Self enumerated questionnaires**

The respondent marks the 'English' box if they identify it as their first language spoken. For respondents whose self identified first language spoken is not English, they put a mark in the 'Other' box.

For people who cannot speak, mark the 'Other' box.

[Previous Page](#)

[Next Page](#)

## **Languages Spoken at Home (LSH)**

[Contents >> Languages Spoken at Home \(LSH\)](#)

This section contains the following subsection :

[Introduction - LSH](#)

[Underlying Concepts - LSH](#)

[Classification and Coding - LSH](#)

[Collection Methods - LSH](#)

[Outputs - LSH](#)

[Measurement Issues - LSH](#)

[Appendix A - Explanatory Script - LSH](#)

[Previous Page](#)

[Next Page](#)

## **Introduction - LSH**

[Contents >> Languages Spoken at Home \(LSH\) >> Introduction - LSH](#)

### **Languages Spoken at Home - Introduction**

---

There is a demand for objective data about languages in Australia. Drivers for these data include government policy on language and literacy, which has evolved rapidly over recent years. Key to the policy are English literacy and the economic relevance of language skills, combined with the significance of language to multiculturalism and diversity within Australia. The goals of Australia's language policy emphasise:

- the importance for all Australians to maintain an appropriate level of written and spoken English
- the importance of Australians learning languages other than English
- the maintenance and learning of Aboriginal and Torres Strait Islander languages reinforced by the Council of Australian Governments' (COAG) policy 'Closing the Gap' and
- the provision of translating and interpreting services.

Stakeholders for language data include health and education service providers, data collection agencies, government agencies, community groups, academics and private sector organisations. Stakeholder interests relate to methodologies for collecting and classifying language information.

To meet these data needs, the Australian Bureau of Statistics (ABS) has incorporated language questions into a range of social statistical collections. The ABS has developed standards for a number of language variables for use when collecting language data both within and outside the ABS.

**Languages Spoken at Home** identifies all languages spoken within the home. This variable enables the collection of data on the full range of languages spoken in Australian homes. To ensure that all relevant languages can be measured the variable includes sign languages.

Program providers have identified that disadvantage in access and equity to government and community programs and services may be experienced by people in Australia whose first language is not English. Data relating to Languages Spoken at Home may aid as an indicator of deficiency in English language competency and other factors associated with cultural background.

The variable Languages Spoken at Home is one of the standard set of indicators for use in measuring cultural and linguistic diversity. It can be used in conjunction with others in the set to understand cultural and ethnic diversity within Australia. Its implementation in administrative and service settings enables the provision of relevant data to assess, measure and monitor service needs, including access and equity requirements.

[Previous Page](#)

[Next Page](#)

## Underlying Concepts - LSH

[Contents](#) >> [Languages Spoken at Home \(LSH\)](#) >> [Underlying Concepts - LSH](#)

### **Languages Spoken at Home - Underlying Concepts**

---

#### **Name of variable**

The name for the variable is Languages Spoken at Home.

#### **Definition of variable**

##### Nominal definition



Languages Spoken at Home is defined as the language or languages spoken by a person in the home, on a regular basis, to communicate with other residents of the home and regular visitors to the home.

Languages Spoken at Home is an attribute of the counting unit 'person'.

### Operational definition

Operationally, Languages Spoken at Home is defined as the language or languages reported by a person as being spoken in the home. There is no restriction on the number of languages reported by the respondent as being spoken in the home.

The definition of language is provided in the Australian Standard Classification of Languages (ASCL), 2011(ABS cat. no. 1267.0)

### **Discussion of conceptual issues**

Languages Spoken at Home identifies all languages spoken within the home. This variable provides data on the range of languages spoken in Australian homes.

Languages Spoken at Home is one of five language variables. The other language variables are First Language Spoken, Main Language Spoken at Home, Main Language Other than English Spoken at Home and Proficiency in Spoken English.

The 'Languages Spoken at Home' variable and the ASCL recognise that approximately one percent of the Australian population use non verbal forms of communication. For coding purposes Auslan and similar sign languages are recognised as separate languages. However Signed English/finger spelling is considered to be another form of English and is coded against English.

Languages Spoken at Home should be used in preference to other language variables when the aim is to collect data on the full range of languages used in the home.

The variable has some limitations as it may not reflect complete language use, for example

- if a particular language is spoken in the home but another language is spoken outside the home, within a person's ethnic community group or
- the variable does not collect data on the frequency that any language is spoken.

[Previous Page](#)

[Next Page](#)

## **Classification and Coding - LSH**

[Contents](#) >> [Languages Spoken at Home \(LSH\)](#) >> [Classification and Coding - LSH](#)

### **Languages Spoken at Home - Classification and Coding**

---

## **The standard classification criteria and structure**

Language data in Australia should be collected, aggregated and disseminated using the Australian Standard Classification of Languages (ASCL), 2011 (ABS cat. no. 1267.0). The two classification criteria used to form categories in the ASCL are:

- genetic affinity, which is the relationship between languages due to evolution from a common ancestral language and
- geographic proximity of language(s) based on the area in which the language originated or first became recognised as a distinct language.

The classification has a three level hierarchical structure, with the exception of the Australian Indigenous Languages broad group, where an extra level has been added between the narrow group and language levels of the classification.

The most detailed level of the classification consists of 432 (216 Indigenous and 216 non Indigenous) base level units. Included in the 432 base level units are 388 specific languages and 44 'not elsewhere classified' (nec) categories, used to code languages not separately listed in the classification.

The second level of the classification comprises 51 narrow groups of languages that are similar in terms of the classification criteria, including seven 'other' categories which consist of languages which do not fit into a particular narrow group.

For three narrow groups of Australian Indigenous languages (Narrow Group 81 Arnhem Land and Daly River Region Languages, Narrow Group 82 Yolngu Matha and Narrow Group 86 Arandic) three digit levels are positioned between the narrow group and language level of the classification. There are 13 three digit level categories. They provide meaningful and useful groups of languages.

The first and most general level of the classification comprises nine broad groups of languages including one 'other' category. Broad groups are formed by aggregating geographically proximate narrow groups.

## **The code structure**

One, two and four digit codes are assigned to the first, second and third level units of the classification respectively.

The first digit identifies the broad group in which each language or narrow group is contained. The first two digits taken together identify the Narrow Group in which each Language is contained.

Within Broad Group 8, Australian Indigenous Languages, there are some narrow groups where an extra level has been added between the narrow group and language levels of the classification. The first three digits taken together identify the additional language groups at this third level.

The four digit codes represent each of the 432 Languages or base level units.

The following examples illustrate the coding scheme;

- a. Broad Group 5 Southern Asian Languages

	Narrow Group	52	Indo-Aryan
	Language	5206	Nepali
b.	Broad Group	8	Australian Indigenous Languages
	Narrow Group	82	Yolngu Matha
	Extra Level	824	Dhuwala
	Language	8245	Madarrpa

### Residual categories and codes

Within each narrow group, a four digit code, consisting of the narrow group code followed by the digits '99', is reserved for a residual 'not elsewhere classified' (nec) or 'other' category. Similarly, within the third level classification of Australian Indigenous Languages the three digit group code may be followed by '9' to denote a 'not elsewhere classified' (nec) or 'other' category. All languages which are not separately identified in the classification are included in these residual 'nec' or 'other' categories of the related classification level.

In each broad group, two digit codes are reserved for residual categories at the narrow group level. These codes consist of the broad group code followed by '9'. These categories are termed 'Other *Broad Group Name*' and consist of separately identified languages which do not fit into any of the narrow groups contained within the broad group, based on the classification criteria.

### Supplementary codes

Supplementary codes are not part of the classification structure. They exist for operational reasons only, and no data would be coded to them if sufficiently detailed responses were obtained in all instances. They are used to process inadequately described responses in statistical collections. The codes are of two types:

- four digit codes ending with one, two or three zeros
- four digit codes commencing with three zeros (operational codes).

Codes ending in zero are described as 'not further defined' (nfd) codes. These codes classify responses to a question about language which cannot be coded to the language level of the classification but which can be coded to a higher level of the classification structure.

Responses which do not relate directly to a particular language category, but which are within the range of languages relating to a particular narrow group, are coded to that narrow group. Such responses are allocated a 'nfd' code consisting of the two digit code of the narrow group followed by 00.

Language responses which do not directly relate to a particular narrow group or language category, but are within the range of languages relating to a particular broad group, are coded to that broad group. These responses are allocated an 'nfd' code consisting of the one digit code of the broad group followed by 000. Language responses which can only be coded at the broad or narrow group levels of the classification can be processed within a collection coded at the four digit level.

Four digit codes commencing with 000 are supplementary codes included for operational purposes to facilitate the coding of responses such as inadequately described languages, etc., which contain insufficient information to be allocated a language, narrow group or

broad group code.

### **Scope of the variable**

The variable **Languages Spoken at Home** applies to all persons.

### **Application of the classification to other variables**

The ASCL can be used for a variety of variables. These include: Main Language Spoken at Home, Main Language Other Than English Spoken at Home, First Language Spoken, Language of Greatest Competency, and Preferred Language.

### **Coding procedures and coding Index**

Language responses to the **Languages Spoken at Home** question are coded to the ASCL, or to one of the supplementary codes, using the guidelines detailed in that classification.

A coding index has been developed to assist in the implementation and use of the ASCL. It contains a comprehensive list of probable responses to questions relating to language and their correct classification codes. Each language response is matched with an entry in the ASCL Coding Index to determine the correct code. Use of the coding index enables responses to be coded accurately to the appropriate category of the classification.

Further information about the classification criteria and coding of data about languages can be found in the ASCL.

Copies of the [Coding Indexes](#) can be found in the ASCL data cube on the ABS website ([www.abs.gov.au](http://www.abs.gov.au)).

[Previous Page](#)

[Next Page](#)

## **Collection Methods - LSH**

[Contents](#) >> [Languages Spoken at Home \(LSH\)](#) >> [Collection Methods - LSH](#)

### **Languages Spoken at Home - Collection Methods**

---

#### **Standard question module**

The standard question module for Languages Spoken at Home elicits the most detailed responses and data.

There are two alternatives which may be used, depending on space and cost considerations:

- Alternative one: a set of 'most likely' languages with tick boxes and an 'Other -please specify' response category for writing languages not specified in the tick box list
- Alternative two: includes adequate response space to enable respondents to write in all languages that apply

Alternative two is more compact but will incur additional coding costs.

### Alternative one

The standard question module for Languages Spoken at Home follows:

**Q1. Which language or languages [do you] [does the person] [does (name)] [will (name of child under two years)] speak at home?**

(Please indicate all languages spoken)

English	<input type="checkbox"/>
Mandarin	<input type="checkbox"/>
Italian	<input type="checkbox"/>
Arabic	<input type="checkbox"/>
Cantonese	<input type="checkbox"/>
Greek	<input type="checkbox"/>
Vietnamese	<input type="checkbox"/>
Spanish	<input type="checkbox"/>
Hindi	<input type="checkbox"/>
Tagalog	<input type="checkbox"/>
Other - Please Specify.....	

For self enumerated questionnaires, respondents should be instructed to mark all boxes that apply as well as write in all non listed languages that apply.

The tick box list for this question module includes languages based on their statistical significance in the Australian context. The list is designed for efficiency when coding responses. The list is reviewed when data indicate that different languages have been more or less frequently reported in the ABS Census.

The tick box list can be extended where circumstances enable longer lists, consistent with the ASCL, to be displayed, such as in electronic collection drop down lists. Similarly, if there are space constraints on paper forms, the tick box list can be truncated.

The 'Other - please specify' category is included to enable respondents, whose languages are not listed, to write their languages in the the spaces provided.

### Alternative two

This alternative question may be used when detailed language data is required from a survey instrument which has space constraints. The question can be asked using only spaces for respondents to write in their languages. Additional coding costs will be incurred since each response will need to be matched to an entry in the Australian Standard Classification of Languages (ASCL), 2011 (ABS cat. no. 1267.0).

**Q1. Which language or languages [do you] [does the person] [does (name)] [will**

**(name of child under two years) speak at home?**  
(Please specify all languages spoken)

.....  
.....  
  
.....  
.....  
  
.....  
.....  
  
.....  
.....

Each alternative to the question module may be accompanied by a brief explanatory note about why Language data is collected and detailed instructions about how to answer the question. The explanatory note can be included with the chosen question module or in supplementary documentation. The recommended text for the explanatory note is in Appendix A

**Standard input categories**

The standard input categories for Languages Spoken at Home are the same as the categories of the ASCL plus the supplementary codes.

**Outputs - LSH**

Contents >> Languages Spoken at Home (LSH) >> Outputs - LSH

**Languages Spoken at Home - Outputs**

---

**Standard output categories**

**Output categories for detailed data**

The hierarchical structure of the Australian Standard Classification of Languages (ASCL), Second Edition (ABS cat. no. 1267.0) allows users the flexibility to output statistics at the level of the classification which best suits their requirements. Data can be presented at broad group level, narrow group level, or the base (language) level. If necessary, significant languages within a narrow group can be presented separately while the remaining languages within the narrow group are aggregated. This approach can also be applied to highlight significant narrow groups within a broad group.

Regardless of the level of aggregation intended for the dissemination of statistics, data should be captured, classified and stored at the base level of the classification wherever possible. Collecting and storing data at the most detailed level of the classification allows the greatest flexibility for the output of statistics, enables more detailed and complex analyses,

facilitates comparisons with previous data using different classifications, and preserves information that provides maximum flexibility for future use of the data.

However, the constraints affecting each statistical collection or other application, such as problems with confidentiality or standard errors, may affect the collection and/or output of data at the lower levels of the classification.

## **Supporting variables**

The variable Languages Spoken at Home requires no supporting variables.

[Previous Page](#)

[Next Page](#)

## **Measurement Issues - LSH**

[Contents](#) >> [Languages Spoken at Home \(LSH\)](#) >> [Measurement Issues - LSH](#)

### **Languages Spoken at Home - Measurement Issues**

---

#### **Measurement issues**

No specific measurement issues apply to the variable Languages Spoken at Home beyond the normal constraints affecting statistical collections, such as problems with confidentiality or standard errors which may affect the collection or output of data for some of the categories at the lower levels of the classification.

#### **Related classifications**

Not applicable.

#### **Contact information**

The Languages Spoken at Home statistical standard has been produced by the Australian Bureau of Statistics. Enquiries should be directed to [standards@abs.gov.au](mailto:standards@abs.gov.au).

[Previous Page](#)

[Next Page](#)

## **Appendix A - Explanatory Script - LSH**

[Contents](#) >> [Languages Spoken at Home \(LSH\)](#) >> [Appendix A - Explanatory Script - LSH](#)

### **Appendix A - Explanatory Script - Languages Spoken at Home**

---

## Why is this question module asked?

Languages Spoken at Home is useful in the collection of data on a full range of languages spoken in Australian homes. When implemented in administrative and service settings data related to this variable may be used to assess, measure and monitor service needs.

## Question module for Detailed Data

### Alternative One

Which language or languages [do you] [does the person] [does (name)] [will (name of child under two years)] speak at home? (Please specify all languages spoken)

English  
Mandarin  
Italian  
Arabic  
Cantonese  
Greek  
Vietnamese  
Spanish  
Hindi  
Tagalog  
Other - please specify .....

### Interviewer based questionnaires

If a respondent identifies that they speak more than one language at home, the interviewer marks all the relevant tick boxes and writes down any language not in the tick box list in the 'Other - please specify' category.

### Self enumerated questionnaires

If a respondent identifies that they speak more than one language at home, they mark all the relevant tick boxes and/or write down any language not in the tick box list in the 'Other - please specify' category. If a respondent's language is not a tick box option, they write the language or languages in the 'Other - please specify' category.

### Alternative Two

Which language or languages [do you] [does the person] [does (name)] [will (name of child under two years)] speak at home? (Please specify all languages spoken)

.....  
.....  
.....  
.....  
.....

### Interviewer based questionnaires

An interviewer writes down all the languages the respondent identifies as speaking at home, in the space provided.

### Self enumerated questionnaires



The respondent writes down all the language they identify as speaking at home, in the space provided.

[Previous Page](#)

[Next Page](#)

## **Main Language Other Than English Spoken at Home (MLOTESH)**

[Contents >> Main Language Other Than English Spoken at Home \(MLOTESH\)](#)

This section contains the following subsection :

- Introduction - MLOTESH
- Underlying Concepts - MLOTESH
- Classification and Coding - MLOTESH
- Collection Methods - MLOTESH
- Outputs - MLOTESH
- Measurement Issues - MLOTESH
- Appendix A - Explanatory Script - MLOTESH

[Previous Page](#)

[Next Page](#)

## **Introduction - MLOTESH**

[Contents >> Main Language Other Than English Spoken at Home \(MLOTESH\) >> Introduction - MLOTESH](#)

### **Main Language Other than English Spoken at Home - Introduction**

---

There is a demand for objective data about languages in Australia. Drivers for these data include government policy on language and literacy, which has evolved rapidly over recent years. Key to the policy are English literacy and the economic relevance of language skills, combined with the significance of language to multiculturalism and diversity within Australia. The goals of Australia's language policy emphasise:

- the importance for all Australians to maintain an appropriate level of written and spoken English
- the importance of Australians learning languages other than English
- the maintenance and learning of Aboriginal and Torres Strait Islander languages reinforced by the Council of Australian Governments' (COAG) policy 'Closing the Gap' and
- the provision of translating and interpreting services.

Stakeholders for language data include health and education service providers, data collection agencies, government agencies, community groups, academics and private sector organisations. Stakeholder interests relate to methodologies for collecting and classifying language information.

To meet these data needs, the Australian Bureau of Statistics (ABS) has incorporated

language questions into a range of social statistical collections. The ABS has developed standards for a number of language variables for use when collecting language data both within and outside the ABS.

**Main Language Other Than English Spoken at Home** identifies the main language, other than English, which are spoken within the home. To ensure that all relevant languages can be measured, the variable includes sign languages.

Program providers have identified that disadvantage in access and equity to government and community programs and services may be experienced by people in Australia whose first language is not English. Data relating to Main Language Other Than English Spoken at Home may aid as an indicator of deficiency in English language competency and other factors associated with cultural background.

The variable Main Language Other Than English Spoken at Home is one of the Minimum Core Indicators from the standard set of indicators for use in measuring cultural and linguistic diversity. It can be used in conjunction with others in the set to understand cultural and ethnic diversity within Australia. Its implementation in administrative and service settings enables the provision of relevant data to assess, measure and monitor service needs, including access and equity requirements.

[Previous Page](#)

[Next Page](#)

## Underlying Concepts - MLOTESH

[Contents >> Main Language Other Than English Spoken at Home \(MLOTESH\) >> Underlying Concepts - MLOTESH](#)

### **Main Language Other than English Spoken at Home - Underlying Concepts**

---

#### **Underlying Concepts**

##### **Name of variable**

The name for the variable is Main Language Other Than English Spoken at Home.

##### **Definition of variable**

###### Nominal definition

Main Language Other Than English Spoken at Home is the main language, other than English, spoken by a person in the home, on a regular basis, to communicate with other residents and regular visitors to the home.

Main Language Other Than English Spoken at Home is an attribute of the counting unit 'person'.

###### Operational definition

Operationally, Main Language Other Than English Spoken at Home is defined as the main language, other than English, reported by a person as being spoken in the home. If the person speaks more than one language at home (not including English), they are asked to report the language spoken most often.

The definition of language is provided in the Australian Standard Classification of Languages (ASCL), 2011 (ABS cat. no. 1267.0)

### **Discussion of conceptual issues**

Main Language Other Than English Spoken at Home identifies languages other than English that people speak within the home environment and in family interactions. If more than one language is spoken in the home, respondents are asked to report the language spoken most often.

This variable provides detailed data about the language background of the Australian population where the main language spoken outside the home is English but a language other than English is mainly spoken at home. It can also assist in measuring possible disadvantage experienced by persons whose usual language is not English.

When used in conjunction with other variables, such as County of Birth, the combined data may identify language usage in first and second generation Australians.

The 'Main Language Other Than English Spoken at Home' variable and the ASCL recognise that approximately one percent of the Australian population use non verbal forms of communication. For coding purposes Auslan and similar sign languages are recognised as separate languages. However Signed English/finger spelling is considered to be another form of English and is coded against English.

The variable has some limitations as it may not reflect complete language use, for example:

- this measure may record those people who mainly speak English at home, but outside the home the person may speak the language of their ethnic community
- the variable may record those people who mainly speak English at home, but who have learnt another language which is occasionally, but not normally, spoken at home.

Main Language Other Than English Spoken at Home is one of five language variables. The other language variables are First Language Spoken, Main Language Spoken at Home, Languages Spoken at Home and Proficiency in Spoken English.

[Previous Page](#)

[Next Page](#)

## **Classification and Coding - MLOTESH**

[Contents >> Main Language Other Than English Spoken at Home \(MLOTESH\) >> Classification and Coding - MLOTESH](#)

### **Main Language Other than English Spoken at Home - Classification and**

### The standard classification criteria and structure

Language data in Australia should be collected, aggregated and disseminated using the Australian Standard Classification of Languages (ASCL), 2011 (ABS cat. no. 1267.0). The two classification criteria used to form categories in the ASCL are:

- genetic affinity, which is the relationship between languages due to evolution from a common ancestral language and
- geographic proximity of language(s) based on the area in which the language originated or first became recognised as a distinct language.

The classification has a three level hierarchical structure, with the exception of the Australian Indigenous Languages broad group, where an extra level has been added between the narrow group and language levels of the classification.

The most detailed level of the classification consists of 432 (216 Indigenous and 216 non Indigenous) base level units. Included in the 432 base level units are 388 specific languages and 44 'not elsewhere classified' (nec) categories, used to code languages not separately listed in the classification.

The second level of the classification comprises 51 narrow groups of languages that are similar in terms of the classification criteria, including seven 'other' categories which consist of languages which do not fit into a particular narrow group.

For three narrow groups of Australian Indigenous languages (Narrow Group 81 Arnhem Land and Daly River Region Languages, Narrow Group 82 Yolngu Matha and Narrow Group 86 Arandic) three digit levels are positioned between the narrow group and language level of the classification. There are 13 three digit level categories. They provide meaningful and useful groups of languages.

The first and most general level of the classification comprises nine broad groups of languages including one 'other' category. Broad groups are formed by aggregating geographically proximate narrow groups.

### The code structure

One, two and four digit codes are assigned to the first, second and third level units of the classification respectively.

The first digit identifies the broad group in which each language or narrow group is contained. The first two digits taken together identify the Narrow Group in which each Language is contained.

Within Broad Group 8, Australian Indigenous Languages, there are some narrow groups where an extra level has been added between the narrow group and language levels of the classification. The first three digits taken together identify the additional language groups at this third level.

The four digit codes represent each of the 432 Languages or base level units.

The following examples illustrate the coding scheme:

- |    |              |      |                                 |
|----|--------------|------|---------------------------------|
| a. | Broad Group  | 5    | Southern Asian Languages        |
|    | Narrow Group | 52   | Indo-Aryan                      |
|    | Language     | 5206 | Nepali                          |
|    |              |      |                                 |
| b. | Broad Group  | 8    | Australian Indigenous Languages |
|    | Narrow Group | 82   | Yolngu Matha                    |
|    | Extra Level  | 824  | Dhuwala                         |
|    | Language     | 8245 | Madarrpa                        |

### **Residual categories and codes**

Within each narrow group, a four digit code, consisting of the narrow group code followed by the digits '99', is reserved for a residual 'not elsewhere classified' (nec) or 'other' category. Similarly, within the third level classification of Australian Indigenous Languages the three digit group code may be followed by '9' to denote a 'not elsewhere classified' (nec) or 'other' category. All languages which are not separately identified in the classification are included in these residual 'nec' or 'other' categories of the related classification level.

In each broad group, two digit codes are reserved for residual categories at the narrow group level. These codes consist of the broad group code followed by '9'. These categories are termed 'Other *Broad Group Name*' and consist of separately identified languages which do not fit into any of the narrow groups contained within the broad group, based on the classification criteria.

### **Supplementary codes**

Supplementary codes are not part of the classification structure. They exist for operational reasons only, and no data would be coded to them if sufficiently detailed responses were obtained in all instances. They are used to process inadequately described responses in statistical collections. The codes are of two types:

- four digit codes ending with one, two or three zeros
- four digit codes commencing with three zeros (operational codes).

Codes ending in zero are described as 'not further defined' (nfd) codes. These codes classify responses to a question about language which cannot be coded to the language level of the classification but which can be coded to a higher level of the classification structure.

Responses which do not relate directly to a particular language category, but which are within the range of languages relating to a particular narrow group, are coded to that narrow group. Such responses are allocated an 'nfd' code consisting of the two digit code of the narrow group followed by 00.

Language responses which do not directly relate to a particular narrow group or language category, but are within the range of languages relating to a particular broad group, are coded to that broad group. These responses are allocated an 'nfd' code consisting of the one digit code of the broad group followed by 000. Language responses which can only be coded at the broad or narrow group levels of the classification can be processed within a collection coded at the four digit level.

Four digit codes commencing with 000 are supplementary codes included for operational purposes to facilitate the coding of responses such as inadequately described languages, etc., which contain insufficient information to be allocated a language, narrow group or broad group code.

### **Scope of the variable**

The variable Main Language Other than English Spoken at Home applies to all persons.

### **Application of the classification to other variables**

The ASCL can be used for a variety of variables. These include: First Language Spoken, Main Language Spoken at Home, Languages Spoken at Home, Language of Greatest Competency, and Preferred Language.

### **Coding procedures and coding Index**

Language responses to the Main Language Other Than English Spoken at Home question are coded to the ASCL, or to one of the supplementary codes, using the guidelines detailed in that classification.

A coding index has been developed to assist in the implementation and use of the ASCL. It contains a comprehensive list of the most probable responses to questions relating to language and their correct classification codes. Each language response is matched with an entry in the ASCL Coding Index to determine the correct code. Use of the coding index enables responses to be coded accurately to the appropriate category of the classification.

Further information about the classification criteria and coding of data about languages can be found in the ASCL.

Copies of the [Coding Indexes](#) can be found in the ASCL data cube on the ABS website ([www.abs.gov.au](http://www.abs.gov.au)).

[Previous Page](#)

[Next Page](#)

## **Collection Methods - MLOTESH**

[Contents >> Main Language Other Than English Spoken at Home \(MLOTESH\) >>](#)  
[Collection Methods - MLOTESH](#)

### **Main Language Other than English Spoken at Home - Collection Methods**

---

#### **Standard question modules**

There are two standard question modules for Main Language Other Than English Spoken at Home:

- The question module for Detailed Data elicits the most detailed responses and data on Main Language Other Than English Spoken at Home.

There are two alternatives for the question module for Detailed Data which may be used, depending on space and cost considerations:

- Alternative one: a set of 'most likely' languages with tick boxes and a 'Yes, Other - please specify' response category for writing a language not specified in the tick box list;
- Alternative two: a tick box for 'No, English Only' and a 'Yes, Other - please specify' write-in category.

Alternative two is more compact but will incur additional coding costs.

- The question module for Minimum Data contains only the tick box options 'No, English only' and 'Yes, Other' and therefore does not collect details as to which language other than English is spoken.

## Question module for Detailed Data

### *Alternative one*

This standard question module for Detailed Data for Main Language Other Than English Spoken at Home is recommended for all collections designed to collect extensive data on main language other than English spoken at home.

The standard question module for obtaining Detailed Data for the variable Main Language Other Than English Spoken at Home is as follows:

**Q1. [[Do you] [Does the person] [Does (name)] [Will (name of child under two years)] speak a language other than English at home?** (If more than one language, indicate the one that is spoken most often.)

<b>No,</b>	English	q
<b>Yes,</b>	Mandarin	q
<b>Yes,</b>	Italian	q
<b>Yes,</b>	Arabic	q
<b>Yes,</b>	Cantonese	q
<b>Yes,</b>	Greek	q
<b>Yes,</b>	Vietnamese	q
<b>Yes,</b>	Spanish	q
<b>Yes,</b>	Hindi	q
<b>Yes,</b>	Tagalog	q
<b>Yes,</b>	Other - Please Specify.....	

For self enumerated questionnaires, respondents should be instructed to mark one box only.

The tick box list for this question module includes languages based on their statistical significance in the Australian context. The list is designed for efficiency when coding responses. The list is reviewed when data indicate that different languages have been more or less frequently reported in the ABS Census.

The tick box list can be extended where circumstances enable longer lists, consistent with

the ASCL, to be displayed, such as in electronic collection drop down lists. Similarly, if there are space constraints on paper forms, the tick box list can be truncated.

The 'Yes, Other' category is included for those people who mainly speak a language in the home that is not offered in the list as a response to the question. 'Please specify' is added to the 'Other' category and a space is provided for respondents to write in their language.

### **Alternative two**

This alternative question may be used when detailed language data is required from a survey instrument which has space constraints. The question uses only the tick box for 'No, English only' and the 'Yes, Other - please specify' option for respondents to write in their language. Additional coding costs will be incurred since each response will need to be matched to an entry in the Australian Standard Classification of Languages (ASCL), 2011 (ABS cat. no. 1267.0).

**Q1. [Do you] [Does the person] [Does (name)] [Will (name of child under two years)] speak a language other than English at home?**

No,	English only	q
Yes,	Other - please specify.....	

### **Question module for Minimum Data**

The standard question module for Minimum Data for the variable Main Language Other than English Spoken at Home is as follows:

**Q1. [Do you] [Does the person] [Does (name)] [Will (name of child under two years)] speak a language other than English at home?**

No,	English only	q
Yes,	Other	q

This Minimum Data question is suitable for collections where there is no requirement for detailed language data. It can be used to identify whether assistance in accessing services and information, due to inadequate English language skills, is required.

Each alternative to the question module may be accompanied by a brief explanatory note about why language data is collected and detailed instructions about how to answer the question. The explanatory note can be included with the chosen question module or in supplementary documentation. The recommended text for the explanatory note is in Appendix A.

### **Standard input categories**

The standard input categories for Main Language Other Than English Spoken at Home are the same as the categories of the ASCL plus the supplementary codes.

### **Input categories for Minimum Data**

The standard input categories for the question module for Minimum Data are:



- English only
- Other language.

[Previous Page](#)

[Next Page](#)

## Outputs - MLOTESH

[Contents](#) >> [Main Language Other Than English Spoken at Home \(MLOTESH\)](#) >> [Outputs - MLOTESH](#)

### Main Language Other than English Spoken at Home - Outputs

---

#### Standard output categories

##### Output categories for Detailed Data

The hierarchical structure of the Australian Standard Classification of Languages (ASCL), Second Edition (ABS cat. no. 1267.0) allows users the flexibility to output statistics at the level of the classification which best suits their requirements. Data can be presented at broad group level, narrow group level, or the base (language) level. If necessary, significant languages within a narrow group can be presented separately while the remaining languages within the narrow group are aggregated. This approach can also be applied to highlight significant narrow groups within a broad group.

Regardless of the level of aggregation intended for the dissemination of statistics, data should be captured, classified and stored at the base level of the classification wherever possible. Collecting and storing data at the most detailed level of the classification allows the greatest flexibility for the output of statistics, enables more detailed and complex analyses, facilitates comparisons with previous data using different classifications, and preserves information that provides maximum flexibility for future use of the data.

However, the constraints affecting each statistical collection or other application, such as problems with confidentiality or standard errors, may affect the collection and/or output of data at the lower levels of the classification.

##### Output categories for Minimum Data

The standard output categories for the question module for Minimum Data are:

- English only
- Other language.

#### Supporting variables

The variable Main Language Other Than English Spoken at Home requires no supporting variables.

## Measurement Issues - MLOTESH

[Contents >> Main Language Other Than English Spoken at Home \(MLOTESH\) >> Measurement Issues - MLOTESH](#)

### Main Language Other than English Spoken at Home - Measurement Issues

---

#### Measurement issues

No specific measurement issues apply to the variable Main Language Other Than English Spoken at Home beyond the normal constraints affecting statistical collections, such as problems with confidentiality or standard errors which may affect the collection or output of data for some of the categories at the lower levels of the classification.

#### Related classifications

Not applicable.

#### Contact information

The Main Language Other than English Spoken at Home statistical standard has been produced by the Australian Bureau of Statistics. Enquiries should be directed to [standards@abs.gov.au](mailto:standards@abs.gov.au).

## Appendix A - Explanatory Script - MLOTESH

[Contents >> Main Language Other Than English Spoken at Home \(MLOTESH\) >> Appendix A - Explanatory Script - MLOTESH](#)

### Appendix A - Explanatory Script - Main Language Other than English Spoken at Home

---

#### Why are these question modules asked?

Data related to Main Language Other than English Spoken at Home can be used to provide information about Australia's language background. It also assists in the planning and implementation of English teaching programs, translation and interpreter services.

#### Question module for Detailed Data

## **Alternative One**

[Do you] [Does the person] [Does (name)] [Will (name of child under two years)] speak a language other than English at Home?

No, English only  
Yes, Mandarin  
Yes, Italian  
Yes, Arabic  
Yes, Cantonese  
Yes, Greek  
Yes, Vietnamese  
Yes, Spanish  
Yes, Hindi  
Yes, Tagalog  
Yes, Other - please specify .....

### **Interview based questionnaires**

If a respondent identifies that they only speak English at home, the interviewer marks the 'No, English only' tick box. When a respondent identifies that they speak another language other than English at home, the interviewer marks the tick box for that language. For a respondent who identifies a language not in the tick box list, the interviewer writes the language name in the 'Yes, Other - please specify' category.

### **Self enumerated questionnaires**

A respondent marks the 'No, English only' tick box, if only English is spoken at home. When a respondent identifies that they speak another language other than English at home, they mark the tick box for that language. With a language not in the tick box list, the respondent writes the language name in the 'Yes, Other - please specify' category.

## **Alternative Two**

[Do you] [Does the person] [Does (name)] [Will (name of child under two years)] speak a language other than English at home?

No, English only  
Yes, Other - please specify .....

### **Interview based questionnaires**

If the respondent only speaks English at home, the interviewer marks the 'No, English only' tick box. For a respondent who identifies a language other than English, the interviewer writes the language name in the 'Yes, Other - please specify' category.

### **Self enumerated questionnaires**

If only English is spoken at home, a respondent marks the 'No, English only' tick box. If a respondent speaks another language other than English at home, they write the language in the 'Yes, Other - please specify category'.

## **Question module for Minimum Data**

[Do you] [Does the person] [Does (name)] [Will (name of child under two years)] speak a

language other than English at home?

No, English only  
Yes, Other

### **Interview based questionnaires**

If the respondent only speaks English at home, the interviewer marks the 'English only' tick box. For all other responses in which a respondent speaks another language other than English, the interviewer marks the tick box for 'Yes, Other'.

### **Self enumerated questionnaires**

The respondent marks the 'English only' tick box, if only English is spoken at home. Where a respondent speaks another language other than English the respondent marks the tick box for 'Yes, Other'.

[Previous Page](#)

[Next Page](#)

## **Main Language Spoken at Home (MLSH)**

[Contents >> Main Language Spoken at Home \(MLSH\)](#)

This section contains the following subsection :

- Introduction - MLSH
- Underlying Concepts - MLSH
- Classification and Coding - MLSH
- Collection Methods - MLSH
- Output - MLSH
- Measurement Issues - MLSH
- Appendix A - Explanatory Script - MLSH

[Previous Page](#)

[Next Page](#)

## **Introduction - MLSH**

[Contents >> Main Language Spoken at Home \(MLSH\) >> Introduction - MLSH](#)

### **Main Language Spoken at Home - Introduction**

---

There is a demand for objective data about languages in Australia. Drivers for these data include government policy on language and literacy, which has evolved rapidly over recent years. Key to the policy are English literacy and the economic relevance of language skills, combined with the significance of language to multiculturalism and diversity within Australia. The goals of Australia's language policy emphasise:

- the importance for all Australians to maintain an appropriate level of written and spoken English
- the importance of Australians learning languages other than English

- the maintenance and learning of Aboriginal and Torres Strait Islander languages reinforced by the Council of Australian Governments' (COAG) policy 'Closing the Gap' and
- the provision of translating and interpreting services.

Stakeholders for language data include health and education service providers, data collection agencies, government agencies, community groups, academics and private sector organisations. Stakeholder interests relate to methodologies for collecting and classifying language information.

To meet these data needs, the Australian Bureau of Statistics (ABS) has incorporated language questions into a range of social statistical collections. The ABS has developed standards for a number of language variables for use when collecting language data both within and outside the ABS.

The **Main Language Spoken at Home** variable identifies the language most commonly spoken in the person's home. To ensure that all relevant languages can be measured, the variable includes sign languages.

Program providers have identified that disadvantage in access and equity to government and community programs and services may be experienced by people in Australia whose first language is not English. Data relating to Main Language Spoken at Home may aid as an indicator of deficiency in English language competency and other factors associated with cultural background.

The variable Main Language Spoken at Home is one of the standard set of indicators for use in measuring cultural and linguistic diversity. It can be used in conjunction with others in the set to understand cultural and ethnic diversity within Australia. Its implementation in administrative and service settings enables the provision of relevant data to assess, measure and monitor service needs, including access and equity requirements.

[Previous Page](#)

[Next Page](#)

## Underlying Concepts - MLSH

[Contents](#) >> [Main Language Spoken at Home \(MLSH\)](#) >> [Underlying Concepts - MLSH](#)

### **Main Language Spoken at Home - Underlying Concepts**

---

#### **Name of variable**

The name of the variable is Main Language Spoken at Home.

#### **Definition of variable**

#### **Nominal definition**

Main Language Spoken at Home is defined as the main language spoken by a person in the home on a regular basis, to communicate with other residents and regular visitors to the home.

Main Language Spoken at Home is an attribute of the counting unit 'person'.

### **Operational definition**

Operationally, Main Language Spoken at Home is defined as the main language reported by a person as being spoken in the home. If the person speaks more than one language at home, they are asked to report the language spoken most often.

The definition of language is provided in the Australian Standard Classification of Languages (ASCL), 2011 (ABS cat. no. 1267.0).

### **Discussion of conceptual issues**

The variable provides data about the languages spoken most frequently or commonly in Australian homes. It does not collect data about the full range of languages spoken by a person in the home. The full range of languages spoken in the home can be collected using the variable "Languages Spoken at Home".

Main Language Spoken at Home is one of five language variables. The other language variables are, First Language Spoken, Languages Spoken at Home, Main Language Other Than English Spoken at Home, and Proficiency in Spoken English.

The Main Language Spoken at Home variable and the ASCL recognise that approximately one percent of the Australian population use non verbal forms of communication. For coding purposes, Auslan and similar sign languages are recognised as separate languages. However Signed English/finger spelling is considered to be another form of English and is coded against English.

[Previous Page](#)

[Next Page](#)

## **Classification and Coding - MLSH**

[Contents](#) >> [Main Language Spoken at Home \(MLSH\)](#) >> [Classification and Coding - MLSH](#)

### **Main Language Spoken at Home - Classification and Coding**

---

#### **The standard classification criteria and structure**

Language data in Australia should be collected, aggregated and disseminated using the Australian Standard Classification of Languages (ASCL), 2011 (ABS cat. no. 1267.0).

The two classification criteria used to form categories in the ASCL are:

- genetic affinity, which is the relationship between languages due to evolution from a common ancestral language, and
- geographic proximity of language(s) based on the area in which the language

originated or first became recognised as a distinct language.

The classification has a three level hierarchical structure, with the exception of the Australian Indigenous Languages broad group, where an extra level has been added between the narrow group and language levels of the classification.

The most detailed level of the classification consists of 432 (216 Indigenous and 216 non Indigenous) base level units. Included in the 432 base level units are 388 specific languages and 44 'not elsewhere classified' (nec) categories, used to code languages not separately listed in the classification.

The second level of the classification comprises 51 narrow groups of languages that are similar in terms of the classification criteria, including seven 'other' categories which consist of languages which do not fit into a particular narrow group.

For three narrow groups of Australian Indigenous languages (Narrow Group 81 Arnhem Land and Daly River Region Languages, Narrow Group 82 Yolngu Matha, and Narrow Group 86 Arandic) three digit levels are positioned between the narrow group and language level of the classification. There are 13 three digit level categories. They provide meaningful and useful groups of languages.

The first and most general level of the classification comprises nine broad groups of languages including one 'other' category. Broad groups are formed by aggregating geographically proximate narrow groups.

### **The code structure**

One, two and four digit codes are assigned to the first, second and third level units of the classification respectively.

The first digit identifies the broad group in which each language or narrow group is contained. The first two digits taken together identify the Narrow Group in which each Language is contained.

Within Broad Group 8, Australian Indigenous Languages, there are some narrow groups where an extra level has been added between the narrow group and language levels of the classification. The first three digits taken together identify the additional language groups at this third level.

The four digit codes represent each of the 432 Languages or base level units.

The following examples illustrate the coding scheme:

- |    |              |      |                                 |
|----|--------------|------|---------------------------------|
| a. | Broad Group  | 5    | Southern Asian Languages        |
|    | Narrow Group | 52   | Indo-Aryan                      |
|    | Language     | 5206 | Nepali                          |
| b. | Broad Group  | 8    | Australian Indigenous Languages |
|    | Narrow Group | 82   | Yolngu Matha                    |
|    | Extra Level  | 824  | Dhuwala                         |
|    | Language     | 8245 | Madarrpa                        |

### **Residual categories and codes**

Within each narrow group, a four digit code consisting of the narrow group code followed by the digits '99', is reserved for a residual 'not elsewhere classified' (nec) or 'other' category. Similarly, within the third level classification of Australian Indigenous Languages, the three digit group code may be followed by '9' to denote a 'not elsewhere classified' (nec) or 'other' category. All languages which are not separately identified in the classification are included in these residual 'nec' or 'other' categories of the related classification level.

In each broad group, two digit codes are reserved for residual categories at the narrow group level. These codes consist of the broad group code followed by '9'. These categories are termed 'Other *Broad Group Name*' and consist of separately identified languages which do not fit into any of the narrow groups contained within the broad group, based on the classification criteria.

### **Supplementary codes**

Supplementary codes are not part of the classification structure. They exist for operational reasons only, and no data would be coded to them if sufficiently detailed responses were obtained in all instances. They are used to process inadequately described responses in statistical collections. The codes are of two types:

- four digit codes ending with one, two or three zeros
- four digit codes commencing with three zeros (operational codes).

Codes ending in zero are described as 'not further defined' (nfd) codes. These codes classify responses to a question about language which cannot be coded to the language level of the classification but which can be coded to a higher level of the classification structure.

Responses which do not relate directly to a particular language category, but which are within the range of languages relating to a particular narrow group, are coded to that narrow group. Such responses are allocated an 'nfd' code consisting of the two digit code of the narrow group followed by 00.

Language responses which do not directly relate to a particular narrow group or language category, but are within the range of languages relating to a particular broad group, are coded to that broad group. These responses are allocated an 'nfd' code consisting of the one digit code of the broad group followed by 000. Language responses which can only be coded at the broad or narrow group levels of the classification can be processed within a collection coded at the four digit level.

Four digit codes commencing with 000 are supplementary codes included for operational purposes to facilitate the coding of responses such as inadequately described languages, etc., which contain insufficient information to be allocated a language, narrow group or broad group code.

### **Scope of the variable**

The variable Main Language Spoken at Home applies to all persons.

### **Application of the classification to other variables**

The ASCL can be used for a variety of variables. These include First Language Spoken, Main Language Other Than English Spoken at Home, Languages Spoken at Home,



Language of Greatest Competency, and Preferred Language.

## **Coding procedures and coding Index**

Language responses to the Main Language Spoken at Home question modules are coded to the ASCL, or to one of the supplementary codes, using the guidelines detailed in that classification.

A coding index has been developed to assist in the implementation and use of the ASCL. It contains a comprehensive list of the probable responses to questions relating to language and their correct classification codes. Each language response is matched with an entry in the ASCL Coding Index to determine the correct code. Use of the coding index enables responses to be coded accurately to the appropriate category of the classification.

Further information about the classification criteria and coding of data about languages can be found in the **ASCL**.

Copies of the [Coding Indexes](#) can be found in the ASCL data cube on the ABS website ([www.abs.gov.au](http://www.abs.gov.au)).

[Previous Page](#)

[Next Page](#)

## **Collection Methods - MLSH**

### **Main Language Spoken at Home - Collection Methods**

---

#### **Standard question modules**

There are two standard question modules for Main Language Spoken at Home:

- The question module for Detailed Data elicits the most detailed responses and data on Main Language Spoken at Home.

There are two alternatives for the question module for Detailed Data which may be used, depending on space and cost considerations:

- Alternative one: a set of 'most likely' languages with tick boxes and an 'Other - please specify' response category for writing a language not specified in the tick box list.
- Alternative two: a tick box for 'English' and an 'Other - please specify' write in category.

Alternative two is more compact but will incur additional coding costs.

- The question module for Minimum Data contains only the tick box options 'English' and 'Other' and therefore does not collect details as to which language other than English is spoken.

**Question module for Detailed Data**

***Alternative one***

The standard question module for Main Language Spoken at Home was designed for self enumerated and interview based surveys. It may also be used in administrative collections. Respondents should be instructed to mark one box only (the language spoken most often) when respondents speak multiple languages at home.

The standard question module for obtaining Detailed Data for the variable Main Language Spoken at Home follows:

**Q1. Which language [do you] [does the person] [does (name)] [will (name of child under two years)] mainly speak at home?**

(If more than one language, indicate the one that is spoken most often.)

English	q
Mandarin	q
Italian	q
Arabic	q
Cantonese	q
Greek	q
Vietnamese	q
Spanish	q
Hindi	q
Tagalog	q
Other - please specify	.....

For self enumerated questionnaires, respondents should be instructed to mark one box only or to write in one response where tick boxes don't apply.

The tick box list for this question module includes languages based on their statistical significance in the Australian context. The list is designed for efficiency when coding responses. The list is reviewed when data indicate that different languages have been more or less frequently reported in the ABS Census and surveys.

The tick box list can be extended where circumstances enable longer lists, consistent with the ASCL, to be displayed, such as in electronic collection drop down lists. Similarly, if there are space constraints on paper forms, the tick box list can be truncated.

The 'Other - please specify' category is included to enable respondents whose main language spoken at home is not listed, to write their language in the space provided.

***Alternative two***

This alternative question module may be used when detailed language data is

required from a survey instrument which has space constraints. The question module uses only the tick box for 'English' and the 'Other - please specify' option for respondents to write in their language. Additional coding costs will be incurred since each response will need to be matched to an entry in the Australian Standard Classification of Languages (ASCL), 2011 ([ABS cat. no. 1267.0](#)).

The second alternative standard question module for obtaining Detailed Data for the variable Main Language Spoken at Home follows:

**Q1. Which language [do you] [does the person] [does (name)] [will (name of child under two years)] mainly speak at home?**

English                      ☐  
Other - please specify .....

**Question module for Minimum Data**

The standard question module for Minimum Data for the variable Main Language Spoken at Home follows:

**Q1. Which language [do you] [does the person] [does (name)] [will (name of child under two years)] mainly speak at home?**

Mark one box only.

English                      ☐  
Other                         ☐

This Minimum Data question module is suitable for collections where there is no requirement for detailed language data. It can be used to identify whether assistance in accessing services and information, due to inadequate English language skills, is required.

Each alternative to the question module may be accompanied by a brief explanatory note about why language data is collected and detailed instructions about how to answer the question. The explanatory note can be included with the chosen question module or in supplementary documentation. The recommended text for the explanatory note is in [Appendix A](#).

**Standard input categories**

**Input categories for Detailed Data**

The standard input categories for Detailed Data for Main Language Spoken at Home are the same as the categories of the ASCL plus the supplementary codes.

## **Input categories for Minimum Data**

The standard input categories for the question module for Minimum Data are:

- English
- Other language

[Previous Page](#)

[Next Page](#)

## **Output - MLSH**

[Contents](#) >> [Main Language Spoken at Home \(MLSH\)](#) >> [Output - MLSH](#)

### **Main Language Spoken at Home - Output**

---

#### **Standard output categories**

##### **Output categories for Detailed Data**

The hierarchical structure of the Australian Standard Classification of Languages (ASCL), Second Edition (ABS cat. no. 1267.0) allows users the flexibility to output statistics at the level of the classification which best suits their requirements. Data can be presented at broad group level, narrow group level, or the base (language) level. If necessary, significant languages within a narrow group can be presented separately while the remaining languages within the narrow group are aggregated. This approach can also be applied to highlight significant narrow groups within a broad group.

Regardless of the level of aggregation intended for the dissemination of statistics, data should be captured, classified and stored at the base level of the classification wherever possible. Collecting and storing data at the most detailed level of the classification allows the greatest flexibility for the output of statistics, enables more detailed and complex analyses, facilitates comparisons with previous data using different classifications, and preserves information that provides maximum flexibility for future use of the data.

However, the constraints affecting each statistical collection or other application, such as problems with confidentiality or standard errors, may affect the collection and/or output of data at the lower levels of the classification.

##### **Output categories for Minimum Data**

The standard output categories for the question module for Minimum Data are:

- English
- Other language

## Supporting variables

The variable Main Language Spoken at Home requires no supporting variables.

[Previous Page](#)

[Next Page](#)

# Measurement Issues - MLSH

[Contents](#) >> [Main Language Spoken at Home \(MLSH\)](#) >> [Measurement Issues - MLSH](#)

## Main Language Spoken at Home - Measurement issues & Related Classifications

---

### Measurement issues

No specific measurement issues apply to the variable Main Language Spoken at Home beyond the normal constraints affecting statistical collections, such as problems with confidentiality or standard errors which may affect the collection or output of data for some of the categories at the lower levels of the classification.

### Related classifications

Not applicable.

### Contact information

The Main Language Spoken at Home statistical standard has been produced by the Australian Bureau of Statistics. Enquiries should be directed to [standards@abs.gov.au](mailto:standards@abs.gov.au).

[Previous Page](#)

[Next Page](#)

# Appendix A - Explanatory Script - MLSH

[Contents](#) >> [Main Language Spoken at Home \(MLSH\)](#) >> [Appendix A - Explanatory Script - MLSH](#)

## Appendix A - Explanatory Script - Main Language Spoken at Home

### Why are these question modules asked?

Main Language Spoken at Home can aid as an indicator of language patterns in Australia. The language data related to this variable can provide useful information about the language

development requirements, any deficiency in English and specific needs a person may have when accessing services.

### **Question module for Detailed Data**

#### **Alternative One**

Which language [do you] [does the person] [does (name)] [will (name of child under two years)] mainly speak at home? (If more than one language is spoken, indicate the one that is spoken most often.)

English  
Mandarin  
Italian  
Arabic  
Cantonese  
Greek  
Vietnamese  
Spanish  
Hindi  
Tagalog  
Other - please specify .....

#### **Interview based questionnaires**

The interviewer marks the box for the language the respondent identifies as the main language spoken. If the respondent identifies a language not in the tick box list, the interviewer writes the name of the identified language in the 'Other - please specify' category. The interviewer could use a prompt if a multiple response is given after asking the question.

#### **Self enumerated questionnaires**

The respondent indicates the main language spoken by putting a mark in one box in the tick box list. If the language identified is not on the list, the respondent can write the language in the 'Other - please specify' category.

#### **Alternative Two**

Which language [do you] [does the person] [does (name)] [will (name of child under two years)] mainly speak at home?

English  
Other - please specify .....

#### **Interview based questionnaires**

The interviewer is instructed to mark the 'English' box if the respondent identifies that English is the main language spoken at home. If the respondent identifies a different language the interviewer can write the name of the language in the 'Other - please specify' category. The interviewer can prompt the respondent if they name more than one language as mainly spoken at home.

#### **Self enumerated questionnaires**

The respondent marks the 'English' box if they identify that English is the main language spoken at home. If a different language is identified by them, they write the name of the language in the 'Other - please specify' category.

### **Question module for Minimum Data**

Which language [do you] [does the person] [does (name)] [will (name of child under two years)] mainly speak at home?

English  
Other

### **Interview based questionnaires**

The interviewer marks the 'English' box if the respondent identifies that English is the main language spoken at home. If the respondent identifies a different language the interviewer marks the 'Other' box.

### **Self enumerated questionnaires**

The respondent marks the 'English' box if they identify that English is the main language spoken at home. The respondent marks the 'Other' box if they identify that another language is spoken at home.

[Previous Page](#)

[Next Page](#)

## **Proficiency in Spoken English**

[Contents >> Proficiency in Spoken English](#)

This section contains the following subsection :

- Introduction - PSE
- Underlying Concepts - PSE
- Classification and Coding - PSE
- Collection Methods - PSE
- Output PSE
- Measurement Issues - PSE
- Appendix A - Explanatory Script - PSE
- Appendix B - Explanatory Notes - Migrant English Language Testing

[Previous Page](#)

[Next Page](#)

## **Introduction - PSE**

[Contents >> Proficiency in Spoken English >> Introduction - PSE](#)

### **Proficiency in Spoken English - Introduction**

---

There is a demand for objective data about languages in Australia. Drivers for these data include government policy on language and literacy, which has evolved rapidly over recent years. Key to the policy are English literacy and the economic relevance of language skills, combined with the significance of language to multiculturalism and diversity within Australia. The goals of Australia's language policy emphasise:

- the importance for all Australians to maintain an appropriate level of written and spoken English
- the importance of Australians learning languages other than English
- the maintenance and learning of Aboriginal and Torres Strait Islander languages reinforced by the Council of Australian Governments' (COAG) policy 'Closing the Gap' and
- the provision of translating and interpreting services.

Stakeholders for language data include health and education service providers, data collection agencies, government agencies, community groups, academics and private sector organisations. Stakeholder interests relate to methodologies for collecting and classifying language information.

To meet these data needs, the Australian Bureau of Statistics (ABS) has incorporated language questions into a range of social statistical collections. The ABS has developed standards for a number of language variables for use when collecting language data both within and outside the ABS.

The variable Proficiency in Spoken English is used to assess the ability in spoken English of people:

- whose first language spoken is a language other than English, or
- who speak a language other than English at home.

Program providers have identified that disadvantage in access and equity to government and community programs and services may be experienced by people in Australia whose first language is not English. Data relating to Proficiency in Spoken English may aid as an indicator of deficiency in English language competency associated with cultural background.

The implementation of the Proficiency in Spoken English variable in administrative and service settings enables the provision of relevant data to assess, measure and monitor service needs, including access and equity requirements.

[Previous Page](#)

[Next Page](#)

## Underlying Concepts - PSE

[Contents](#) >> [Proficiency in Spoken English](#) >> [Underlying Concepts - PSE](#)

### Proficiency in Spoken English - Underlying Concepts

---

#### Name of variable

The name of the variable is Proficiency in Spoken English.



## **Nominal definition**

The variable Proficiency in Spoken English is defined as the ability to speak English in every day situations.

Proficiency in Spoken English is an attribute of the counting unit 'person'. That is, it is an attribute of an individual as opposed to a family or household.

## **Operational definition**

Operationally, the variable is defined as the self-assessed level of ability to speak English, asked of people whose first language spoken is a language other than English or who speak a language other than English at home.

The definition of language is provided in the Australian Standard Classification of Languages (ASCL), 2011 (ABS cat. no. 1267.0).

## **Discussion of conceptual issues**

This variable was formerly called 'Proficiency in English', a name which implied a measurement of all aspects of proficiency: speaking, listening, writing and reading. However, in ABS collections this variable has only ever been a measurement of spoken English proficiency and therefore 'Proficiency in Spoken English' is a more accurate label.

English language proficiency can be tested through a range of examination systems. However in the current Australian immigration context, the Department of Immigration and Border Protection (DIBP) (formerly the Department of Immigration and Citizenship) has stipulated minimum levels of English language proficiency which can be equated to particular scores on the International English Language Testing System (IELTS) tests (see Appendix B for more details). IELTS scores are obtained through four test components (speaking, reading, listening and writing) as part of the visa application process. Gaining English language proficiency is regarded as a key step for migrants to take towards successful settlement in Australia.

While social surveys cannot measure proficiency in spoken English with the same degree of precision as formal tests, ABS pilot testing of this question prior to the 1981 Census assessed the degree of correspondence between self assessed proficiency and a formal test of English proficiency. There was an overall correspondence between the numbers with low test ratings and those reporting their ability as 'Not Well' or 'Not at All', indicating that the variable provides a reasonably reliable measure of the number of people who may be in need of targeted services.

In most ABS surveys, a question on Proficiency in Spoken English is asked only of respondents whose first language spoken is a language other than English, or those who speak a language other than English at home. This is because the variable is used primarily to identify people who may experience a disadvantage as a result of a lack of competence in spoken English.

[Previous Page](#)

[Next Page](#)

## **Classification and Coding - PSE**

## Proficiency in Spoken English - Classification and Coding

---

The standard classification is a flat or single-level classification consisting of four categories:

Very well  
Well  
Not well  
Not at all

The code structure is:

1 Very  
well  
2 Well  
3 Not  
well  
4 Not at  
all

### Supplementary codes

Two supplementary codes are reserved for the following conditions:

0 Not stated/Inadequately described

9 Not applicable

The 'Not stated/Inadequately described' supplementary code is used for responses that are not stated or are ambiguous in the context of the standard classification and cannot be accurately coded to any of the four categories in the classification.

The 'Not applicable' supplementary code is used for respondents who did not speak a language other than English for their first language spoken and do not speak a language other than English at home.

### Scope of the variable

The variable Proficiency in Spoken English applies to all persons who did not speak English as their first language or who speak another language at home.

### Application of the classification to other variables

The Proficiency in Spoken English classification is not applicable to any other variable.

### Coding procedures

Proficiency in Spoken English is self-coded to the classification.

## Collection Methods - PSE

Contents >> Proficiency in Spoken English >> Collection Methods - PSE

### Proficiency in Spoken English - Collection Methods

---

The question on Proficiency in Spoken English is asked after a question which identifies people whose first language spoken is a language other than English, or people who speak a language other than English at home. This sequencing is necessary because the Proficiency in Spoken English variable is used to measure the number of people who may potentially suffer a disadvantage as a result of a lack of competence in spoken English, rather than assess the spoken English capacity of the Australian population in general. It is considered unnecessary and inappropriate to ask people whose only language is English, how well they speak it.

The Proficiency in Spoken English question should be asked after a language question which filters out people who only speak English. There are four filter questions to choose from, derived from the following ABS statistical standards:

- Filter 1 - First Language Spoken
- Filter 2 - Main Language Spoken at Home
- Filter 3 - Main Language Other than English Spoken at Home
- Filter 4 - Languages Spoken at Home

#### Filter question modules for standard question

#### Standard question modules

#### Short question modules

The standard and short question modules may be accompanied by a brief explanatory note about why language data is collected and detailed instructions about how to answer the question. The explanatory note can be included with the chosen question module or in supplementary documentation. The recommended text for the explanatory note is in Appendix A.

#### Standard input categories

The standard input categories for Proficiency in Spoken English are the same as the categories of the classification and also includes the two supplementary categories, 'Not stated/Inadequately described' and 'Not applicable':

- Very well
- Well
- Not well
- Not at all
- Not stated/Inadequately described
- Not applicable

This section contains the following subsection :

Collection Methods (PSE) - Filter Questions Modules for Standard Question

Collection Methods (PSE) - Standard Question Modules

Collection Methods (PSE) - Short Question Modules

[Previous Page](#)

[Next Page](#)

## Collection Methods (PSE) - Filter Questions Modules for Standard Question

[Contents](#) >> [Proficiency in Spoken English](#) >> [Collection Methods - PSE](#) >> [Collection Methods \(PSE\) - Filter Questions Modules for Standard Question](#)

### Proficiency in Spoken English - Filter question modules for standard question

---

Only ONE filter question is required. The preferred filter question for this variable is First Language Spoken.

The tick box lists for the filter questions include languages of statistical significance in the Australian context. The list is designed for efficiency when coding responses and is modified periodically, according to the frequency of languages reported in the ABS Census and surveys.

The tick box list can be extended where circumstances enable longer lists, consistent with the ASCL, such as in electronic collection drop down lists. Similarly, if there are space constraints on paper forms, the response options available within the list can be reduced.

The 'Other - please specify' category is included to enable respondents whose language(s) is/are not listed, to write their language(s) in the space provided.

### The Standard Filter Question Modules

#### FILTER 1 - First Language Spoken

**Q1. Which language [did you] [did the person] [did (name)] [will (name of child under two years)] first speak as a child?**

Mark one box only.

English	<input type="checkbox"/>
Mandarin	<input type="checkbox"/>
Italian	<input type="checkbox"/>

Arabic	q
Cantonese	q
Greek	q
Vietnamese	q
Spanish	q
Hindi	q
Tagalog	q
Other - Please Specify.....	

**Sequence guide:**

- If response is 'English', then No More Questions.
- If response is any language other than English, then go to (Q2) - Proficiency in Spoken English question.

**OR**

**FILTER 2 - Main Language Spoken at Home**

**Q1. Which language [do you] [does the person] [does (name)] [will (name of child under two years)] mainly speak at home?** (If more than one language, indicate the one that is spoken most often.)

English	q
Mandarin	q
Italian	q
Arabic	q
Cantonese	q
Greek	q
Vietnamese	q
Spanish	q
Hindi	q
Tagalog	q
Other - Please Specify.....	

**Sequence guide:**

- If response is 'English', then No More Questions.
- If response is any language other than English, then go to (Q2) - Proficiency in Spoken English question.

**OR**

**FILTER 3 - Main Language Other Than English Spoken at Home**

**Q1. [Do you] [Does the person] [Does (name)] [Will (name of child under two years)] speak a language other than English at home?** (If more than one language, indicate the one that is spoken most often.)

<b>No,</b>	English	q
<b>Yes,</b>	Mandarin	q
<b>Yes,</b>	Italian	q
<b>Yes,</b>	Arabic	q
<b>Yes,</b>	Cantonese	q
<b>Yes,</b>	Greek	q

Yes,	Vietnamese	q
Yes,	Spanish	q
Yes,	Hindi	q
Yes,	Tagalog	q
Yes,	Other - Please Specify.....	

#### Sequence guide:

- If response is 'No, English', then No More Questions.
- If response is any language other than English, then go to (Q2) - Proficiency in Spoken English question.

**OR**

#### **FILTER 4 - Languages Spoken at Home**

**Q1. Which language or languages [do you] [does the person] [does (name)] [will (name of child under two years)] speak at home? (Please indicate all languages spoken)**

English	q
Mandarin	q
Italian	q
Arabic	q
Cantonese	q
Greek	q
Vietnamese	q
Spanish	q
Hindi	q
Tagalog	q
Other - Please Specify.....	

#### Sequence guide:

- If the only response is 'English', then No More Questions.
- If the response includes any other language then go to (Q2) - Proficiency in Spoken English question.

[Previous Page](#)

[Next Page](#)

## **Collection Methods (PSE) - Standard Question Modules**

[Contents >> Proficiency in Spoken English >> Collection Methods - PSE >> Collection Methods \(PSE\) - Standard Question Modules](#)

### **Proficiency in Spoken English - Standard question modules**

---

There are two standard questions for Proficiency in Spoken English:

- self-enumerated surveys
- surveys conducted by interview.

## Self-enumerated surveys

### Q2. How well [do you] [does the person] speak English?

Very well

Well

Not well

Not at all

## Interview-based surveys

### Q2. Do you consider [you speak] [(name) speaks] English very well, well, or not well?

Very well

Well

Not well

Not at all

A separate question is required for interview based surveys as, unlike a self enumerated question where respondents can see the response choices and use them to interpret the question, respondents may be unaware of how to rate their proficiency when the interviewer asks the question "How well do you speak English?". Including the options 'very well', 'well' and 'not well' when asking the question gives the respondents some idea of how to rate their proficiency. It is important that this question records the respondent's own perception of how well they rate their English speaking skills.

[Previous Page](#)

[Next Page](#)

## Collection Methods (PSE) - Short Question Modules

[Contents >> Proficiency in Spoken English >> Collection Methods - PSE >> Collection Methods \(PSE\) - Short Question Modules](#)

### Proficiency in Spoken English - Short question modules

---

There are times when a survey needs to collect data about proficiency in spoken English when the additional data related to other language detail does not need to be collected. A minimum data filter question can be used in these circumstances.

The preferred minimum data filter question is drawn from the 'Main Language Other Than English Spoken at Home' variable and is as follows:

### Q1 [Do you] [Does the person] [Does (name)] [Will (name of child under two years)] speak a language other than English at home?

No,	English only	q
Yes,	Other	q

**Sequence guide:**

- If response is 'No, English only', then No More Questions.
- If response is 'Yes, Other', then go to (Q2) - Proficiency in Spoken English question.

The Proficiency in Spoken English question will vary depending on whether the survey is:

- self-enumerated
- conducted by interview.

### **Self-enumerated surveys**

#### **Q2. How well [do you] [does the person] speak English?**

Very well  
Well  
Not well  
Not at all

### **Interview-based surveys**

#### **Q2. Do you consider [you speak] [(name) speaks] English very well, well, or not well?**

Very well  
Well  
Not well  
Not at all

[Previous Page](#)

[Next Page](#)

## **Output PSE**

[Contents](#) >> [Proficiency in Spoken English](#) >> [Output PSE](#)

### **Proficiency in Spoken English - Output**

---

#### **Standard output categories**

The standard output categories are the same as the categories of the classification and includes the supplementary categories if applicable:

- Very well
- Well
- Not well
- Not at all
- Not stated/Inadequately described
- Not applicable

As the distinction between the categories 'Very well' and 'Well' is somewhat subjective, they



can be amalgamated into the output category 'Well or Very well'. Similarly, 'Not well' and 'Not at all' can be amalgamated into the output category 'Not well or Not at all'. In this case, the output categories would be as follows:

- Well or Very well
- Not well or Not at all
- Not stated/Inadequately described
- Not applicable

## **Supporting variables**

The variable Proficiency in Spoken English requires no supporting variables.

[Previous Page](#)

[Next Page](#)

# **Measurement Issues - PSE**

[Contents](#) >> [Proficiency in Spoken English](#) >> [Measurement Issues - PSE](#)

## **Proficiency in Spoken English - Measurement Issues & Related Classifications**

---

### **Measurement Issues**

The ABS considers it is inappropriate to ask people who use sign languages how well they speak English. This should be avoided wherever possible if these people are filtered to the Proficiency in Spoken English question.

### **Related classifications**

There are no related classifications

### **Contact information**

The Proficiency in Spoken English statistical standard has been produced by the Australian Bureau of Statistics. Enquiries should be directed to [standards@abs.gov.au](mailto:standards@abs.gov.au). The ABS Privacy Policy outlines how the ABS will handle any personal information that you provide to us.

[Previous Page](#)

[Next Page](#)

# Appendix A - Explanatory Script - PSE

[Contents >> Proficiency in Spoken English >> Appendix A - Explanatory Script - PSE](#)

## Appendix A - Explanatory Script - Proficiency in Spoken English

---

Australia's main language is English. However, past Censuses have shown that over three million Australians spoke a language other than English at home.

Language questions are asked in Australia to learn which other languages are spoken and how well English is spoken. Data from responses to these questions make it easier to plan for English teaching programs and for translation and interpreter services.

In most ABS surveys, a question on Proficiency in Spoken English is asked only of respondents whose first language spoken was a language other than English, or those who speak a language other than English at home. The question aims to collect the data based on the self assessed view of respondents about how well they can conduct a conversation in English about every day things. There are four levels of self assessed proficiency in spoken English:

Very well

Well

Not well

Not at all

[Previous Page](#)

[Next Page](#)

## Appendix B - Explanatory Notes - Migrant English Language Testing

[Contents >> Proficiency in Spoken English >> Appendix B - Explanatory Notes - Migrant English Language Testing](#)

### Appendix B - Explanatory Notes - Migrant English Language Testing

---

English language proficiency can be tested through a range of examination systems. However in the current Australian immigration context, the Department of Immigration and Border Protection (DIBP) (formerly the Department of Immigration and Citizenship) has stipulated minimum levels of English language proficiency which can be equated to particular scores on the International English Language Testing System (IELTS) tests. IELTS scores are obtained through four test components (speaking, reading, listening and writing) as part of the visa application process. Gaining English language proficiency is regarded as a key step for migrants to take towards successful settlement in Australia.

The following terms for levels of English language proficiency are used by DIBP when assessing eligibility for certain visas:

- **For points based skilled migration\***
  - Competent English - IELTS (or equivalent) score 6 for each of the four test components - minimum requirement no additional points
  - Proficient English - IELTS (or equivalent ) score 7 for each of the four test components - 10 additional points
  - Superior English - IELTS (or equivalent) score 8 for each of the four test components - 20 additional points
- **Business skilled migration depending on type of business visa**
  - Vocational English - IELTS (or equivalent) score 5 for each of the four test components
  - Functional English - IELTS (or equivalent) score 4.5 for each of the four test components
  - Functional English or additional visa fees to meet the costs of the Adult Migrant Education Program (AMEP) classes are required for adult family members of business visa applicants
  - Functional English or additional visa fees to meet the costs of AMEP classes are required for some business visa applicants (e.g. investors)
- **Other migrants**
  - Where adult migrants do not meet the minimum level 'Functional English', additional charges are added to visa applications to cover AMEP.
  - Humanitarian and family reunion visas also provide AMEP tuition to enable migrants to reach the functional English level

\* Certain occupations (e.g. health professionals) are required to meet the IELTS criteria in relevant Occupational English tests.

[Previous Page](#)

## History of Changes

**20/03/2014**

The reviews of the full suite of five ABS Language Standards have now been completed by adding the updated Proficiency in Spoken English Standard on 20 March, 2014.

Key changes to the Proficiency in Spoken English Standard include:

- An additional filter question module which will enable the collection of data about all languages spoken.
- The addition of a short question module where information about which other languages are spoken is not necessary to data being collected.

The main objectives of the reviews of the suite of Language Standards were to:

- maintain relevance
- maintain consistency in the collection of high quality data
- improve data usability
- rectify identified issues

---

© Commonwealth of Australia

All data and other material produced by the Australian Bureau of Statistics (ABS) constitutes Commonwealth copyright administered by the ABS. The ABS reserves the right to set out the terms and conditions for the use of such material. Unless otherwise noted, all material on this website – except the ABS logo, the Commonwealth Coat of Arms, and any material protected by a trade mark – is licensed under a Creative Commons Attribution 2.5 Australia licence